



PCT

## RAW SEQUENCE LISTING

DATE: 03/16/2006

PATENT APPLICATION: US/10/539,450

TIME: 11:40:04

Input Set : N:\SMITH\PTO.TS.txt

Output Set: N:\CRF4\03162006\J539450.raw

3 <110> APPLICANT: Hisashi NARIMATSU  
 4 Akira TOGAYACHI  
 5 Niro INABA  
 6 Toru HIRUMA  
 7 Yasuko ISHIZUKA  
 8 NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND  
 9 TECHNOLOGY  
 11 <120> TITLE OF INVENTION: BETA1,3-N-ACETYL-D-GALACTOSAMINE TRANSFERASE PROTEIN,  
 12 NUCLEIC ACID ENCODING THE SAME AND METHOD OF EXAMINING CANCERATION  
 13 USING THE SAME  
 15 <130> FILE REFERENCE: 159-90 / PC/S-84-6  
 17 <140> CURRENT APPLICATION NUMBER: US 10/539,450  
 C--> 18 <141> CURRENT FILING DATE: 2005-06-20  
 20 <150> PRIOR APPLICATION NUMBER: PCT/JP2004/000608  
 21 <151> PRIOR FILING DATE: 2004-01-23  
 23 <150> PRIOR APPLICATION NUMBER: JP 392555/2003  
 24 <151> PRIOR FILING DATE: 2003-11-21  
 26 <150> PRIOR APPLICATION NUMBER: JP 285310/2003  
 27 <151> PRIOR FILING DATE: 2003-08-01  
 29 <150> PRIOR APPLICATION NUMBER: JP 014792/2003  
 30 <151> PRIOR FILING DATE: 2003-01-23  
 32 <160> NUMBER OF SEQ ID NOS: 27  
 34 <170> SOFTWARE: MS Word  
 36 <210> SEQ ID NO: 1  
 37 <211> LENGTH: 1503  
 38 <212> TYPE: DNA  
 39 <213> ORGANISM: Homo sapiens  
 41 <400> SEQUENCE: 1  
 42 atgcgaaact ggctggtgct gctgtgcccg tgtgtgctcg gggccgcgct gcacctctgg 60  
 43 ctgcggctgc gctccccgcc gcccgctgc gcctccgggg ccggccctgc agatcagttg 120  
 44 gccttatttc ctcaagtggaa atctactcac tatgatgtgg tagttggcgt gttgtcagct 180  
 45 cgcaataacc atgaacttcg aaacgtgata agaagcacct ggatgagaca ttgtctacag 240  
 46 catcccatat taagtcaacg tgtgcttggt aagttcataa taggtgctca tggctgtgaa 300  
 47 gtgcctgtgg aagacaggga agatccttat tcctgtaaag tactcaacat cacaaatcca 360  
 48 gttttgaatc aggaaattga agcgttcagt ctgtccgaag acacttcacg ggggctgcct 420  
 49 gaggatcgag ttgtcagcgt gagtttccga gttctctacc ccatcgttat taccagtctt 480  
 50 ggagtgttct acgatgccaa tgatgtgggt ttccagagga acatcactgt caaactttat 540  
 51 caggcagaac aagaggaggc cctcttcatt gctcgttca gtcctccaag ctgtggtgtg 600  
 52 caggtgaaca agctgtggtg caagcccgtg gaacaattca tcttaccaga gagctttgaa 660  
 53 ggtacaatcg tgtgggagag ccaagacctc cacggccttg tgtcaagaaa tctccacaaa 720  
 54 gtgacagtga atgatggagg gggagttctc agagtcatta cagctgggga ggggtgcattg 780  
 55 cctcatgaat tcttggaagg tgtggaggga gttgcagggt gttttatata tactattcag 840  
 56 gaaggtgatg ctctcttaca caaccttcac tctcgccctc aaagacttat tgatcatata 900

## RAW SEQUENCE LISTING

DATE: 03/16/2006

PATENT APPLICATION: US/10/539,450

TIME: 11:40:04

Input Set : N:\SMITH\PTO.TS.txt

Output Set: N:\CRF4\03162006\J539450.raw

```

57 aggaatctcc atgaggaaga tgccttactg aaggaggaaa gcagcatcta tgatgatatt 960
58 gtttttgtgg atgttgtcga cacttatcgt aatgttcctg caaaattatt gaacttctat 1020
59 agatggactg tggaaacaac gagcttcaat ttgttgctga agacagatga tgactgttac 1080
60 atagacctcg aagctgtatt taataggatt gtccaaaaga atctggatgg gcctaatttt 1140
61 tgggtggggaa atttcagact gaattgggca gttgaccgaa ccggaagtgt gcaggagtgt 1200
62 gagtaccgca gccccgctta ccctgccttt gcatgtgggt caggatatgt gatctccaag 1260
63 gacatcgtca agtggctggc aagcaactcg gggagggttaa agacctatca ggggtgaagat 1320
64 gtaagcatgg gcatctggat ggctgccata ggacctaaaa gataccagga cagtctgtgg 1380
65 ctgtgtgaga agacctgtga gacaggaatg ctgtcttctc ctcagtattc tccgtgggaa 1440
66 ctgacggaac tgtggaaact gaaggaacgg tgcggtgatc cttgtcgatg tcaagcaaga 1500
67 taa 1503

70 <210> SEQ ID NO: 2
71 <211> LENGTH: 500
72 <212> TYPE: PRT
73 <213> ORGANISM: Homo sapiens
75 <400> SEQUENCE: 2
76 Met Arg Asn Trp Leu Val Leu Leu Cys Pro Cys Val Leu Gly Ala Ala
77 1 5 10 15
79 Leu His Leu Trp Leu Arg Leu Arg Ser Pro Pro Pro Ala Cys Ala Ser
80 20 25 30
82 Gly Ala Gly Pro Ala Asp Gln Leu Ala Leu Phe Pro Gln Trp Lys Ser
83 35 40 45
85 Thr His Tyr Asp Val Val Val Gly Val Leu Ser Ala Arg Asn Asn His
86 50 55 60
88 Glu Leu Arg Asn Val Ile Arg Ser Thr Trp Met Arg His Leu Leu Gln
89 65 70 75 80
91 His Pro Thr Leu Ser Gln Arg Val Leu Val Lys Phe Ile Ile Gly Ala
92 85 90 95
94 His Gly Cys Glu Val Pro Val Glu Asp Arg Glu Asp Pro Tyr Ser Cys
95 100 105 110
97 Lys Leu Leu Asn Ile Thr Asn Pro Val Leu Asn Gln Glu Ile Glu Ala
98 115 120 125
100 Phe Ser Leu Ser Glu Asp Thr Ser Ser Gly Leu Pro Glu Asp Arg Val
101 130 135 140
103 Val Ser Val Ser Phe Arg Val Leu Tyr Pro Ile Val Ile Thr Ser Leu
104 145 150 155 160
106 Gly Val Phe Tyr Asp Ala Asn Asp Val Gly Phe Gln Arg Asn Ile Thr
107 165 170 175
109 Val Lys Leu Tyr Gln Ala Glu Gln Glu Glu Ala Leu Phe Ile Ala Arg
110 180 185 190
112 Phe Ser Pro Pro Ser Cys Gly Val Gln Val Asn Lys Leu Trp Tyr Lys
113 195 200 205
115 Pro Val Glu Gln Phe Ile Leu Pro Glu Ser Phe Glu Gly Thr Ile Val
116 210 215 220
118 Trp Glu Ser Gln Asp Leu His Gly Leu Val Ser Arg Asn Leu His Lys
119 225 230 235 240
121 Val Thr Val Asn Asp Gly Gly Gly Val Leu Arg Val Ile Thr Ala Gly
122 245 250 255
124 Glu Gly Ala Leu Pro His Glu Phe Leu Glu Gly Val Glu Gly Val Ala

```

## RAW SEQUENCE LISTING

DATE: 03/16/2006

PATENT APPLICATION: US/10/539,450

TIME: 11:40:04

Input Set : N:\SMITH\PTO.TS.txt

Output Set: N:\CRF4\03162006\J539450.raw

```

125          260          265          270
127 Gly Gly Phe Ile Tyr Thr Ile Gln Glu Gly Asp Ala Leu Leu His Asn
128          275          280          285
130 Leu His Ser Arg Pro Gln Arg Leu Ile Asp His Ile Arg Asn Leu His
131          290          295          300
133 Glu Glu Asp Ala Leu Leu Lys Glu Glu Ser Ser Ile Tyr Asp Asp Ile
134 305          310          315          320
136 Val Phe Val Asp Val Val Asp Thr Tyr Arg Asn Val Pro Ala Lys Leu
137          325          330          335
139 Leu Asn Phe Tyr Arg Trp Thr Val Glu Thr Thr Ser Phe Asn Leu Leu
140          340          345          350
142 Leu Lys Thr Asp Asp Asp Cys Tyr Ile Asp Leu Glu Ala Val Phe Asn
143          355          360          365
145 Arg Ile Val Gln Lys Asn Leu Asp Gly Pro Asn Phe Trp Trp Gly Asn
146          370          375          380
148 Phe Arg Leu Asn Trp Ala Val Asp Arg Thr Gly Lys Trp Gln Glu Leu
149 385          390          395          400
151 Glu Tyr Pro Ser Pro Ala Tyr Pro Ala Phe Ala Cys Gly Ser Gly Tyr
152          405          410          415
154 Val Ile Ser Lys Asp Ile Val Lys Trp Leu Ala Ser Asn Ser Gly Arg
155          420          425          430
157 Leu Lys Thr Tyr Gln Gly Glu Asp Val Ser Met Gly Ile Trp Met Ala
158          435          440          445
160 Ala Ile Gly Pro Lys Arg Tyr Gln Asp Ser Leu Trp Leu Cys Glu Lys
161          450          455          460
163 Thr Cys Glu Thr Gly Met Leu Ser Ser Pro Gln Tyr Ser Pro Trp Glu
164 465          470          475          480
166 Leu Thr Glu Leu Trp Lys Leu Lys Glu Arg Cys Gly Asp Pro Cys Arg
167          485          490          495
169 Cys Gln Ala Arg
170          500
173 <210> SEQ ID NO: 3
174 <211> LENGTH: 1515
175 <212> TYPE: DNA
176 <213> ORGANISM: Mouse
178 <400> SEQUENCE: 3
179 atgcgaaact ggctggtgct gctgtgccct tgcgtgctcg gggccgcgct gcacctctgg 60
180 cacctctggc tccgttcccc gccggacccc cacaacaccg ggcccagcgc ggcagatcaa 120
181 tcagccttat ttctcactg gaaatttagc cactatgatg tggtagttgg tgtgttatca 180
182 gctcgaaata accacgaact tcgaaatgtg ataaggaaca cctggctgaa gaatttgctg 240
183 catcatccta cattaagtca acgtgtgctt gtgaagttca taataggtgc ccgtggctgt 300
184 gaagtgcctg tggaagacag ggaggatcct tactcctgcc gactgctcaa catcaccaat 360
185 ccagttttga atcaagaaat tgaggcattc agctttcctg aagatgcctc ctcatctaga 420
186 ctctctgaag accgagttgt cagcgtgagc ttcagagttc tctaccaat cgtgattacc 480
187 agtcttgag tgttctacga tgccagtgat gttggttttc aaaggaacat cacagtcaag 540
188 ttgtatcaga cagagcagga ggaggccctt ttcatcgccc gattcagtc tccaagttgt 600
189 ggcgtacaag tgaacaagct ctggtataag cccgtggaac agttcatctt accagagagc 660
190 tttgaaggta caatcgtgtg ggaaagccaa gatctccatg gcctcgtgtc cagaaacctg 720
191 cacagagtga cagtgaatga tggagggggg gttctcagag tccttgacag tggggaaggg 780

```

## RAW SEQUENCE LISTING

DATE: 03/16/2006

PATENT APPLICATION: US/10/539,450

TIME: 11:40:04

Input Set : N:\SMITH\PTO.TS.txt

Output Set: N:\CRF4\03162006\J539450.raw

```

192 gcactgcctc atgaattcat ggaagggtgtg gagggagttg cgggtggcctt tatctacact 840
193 gttcaggaag gtgatgcact attaagaagc ctttattctc ggccccagag acttgcagat 900
194 cacatacagg atctgcaggt ggaagatgcc ttactgcagg aggaaagcag tgtccatgac 960
195 gacattgtct tcgtggatgt tgtggatact taccggaatg ttcttgcaaa attactgaac 1020
196 ttctatagat ggactgtgga atccaccagc ttcgatttgc tgctcaagac agatgacgac 1080
197 tgttatatag acttagaagc tgtgtttaat agaattgctc agaagaatct agatgggcct 1140
198 aatttttggt ggggaaattt cagggttgaat tgggcagtggt acagaaccgg aaaatggcag 1200
199 gagctggaat acccgagccc ggcttaccct gcctttgcat gtgggtcagg gtatgtgatc 1260
200 tccaaggata tcgttgactg gctggcaggc aactccagaa ggttaaagac ctatcagggt 1320
201 gaagatgtca gcatgggcat ttggatggca gccataggac ctaaaagaca ccaggacagc 1380
202 ctgtggctgt gtgagaaaac ctgtgagaca ggaatgctgt cttctcctca gtactcacca 1440
203 gaagagctga gcaaaactctg ggaactgaag gagctgtgtg gggatccttg tcagtgtgaa 1500
204 gcaaaagtac gatga 1515
207 <210> SEQ ID NO: 4
208 <211> LENGTH: 504
209 <212> TYPE: PRT
210 <213> ORGANISM: Mouse
212 <400> SEQUENCE: 4
213 Met Arg Asn Trp Leu Val Leu Leu Cys Pro Cys Val Leu Gly Ala Ala
214 1 5 10 15
216 Leu His Leu Trp His Leu Trp Leu Arg Ser Pro Pro Asp Pro His Asn
217 20 25 30
219 Thr Gly Pro Ser Ala Ala Asp Gln Ser Ala Leu Phe Pro His Trp Lys
220 35 40 45
222 Phe Ser His Tyr Asp Val Val Val Gly Val Leu Ser Ala Arg Asn Asn
223 50 55 60
225 His Glu Leu Arg Asn Val Ile Arg Asn Thr Trp Leu Lys Asn Leu Leu
226 65 70 75 80
228 His His Pro Thr Leu Ser Gln Arg Val Leu Val Lys Phe Ile Ile Gly
229 85 90 95
231 Ala Arg Gly Cys Glu Val Pro Val Glu Asp Arg Glu Asp Pro Tyr Ser
232 100 105 110
234 Cys Arg Leu Leu Asn Ile Thr Asn Pro Val Leu Asn Gln Glu Ile Glu
235 115 120 125
237 Ala Phe Ser Phe Pro Glu Asp Ala Ser Ser Ser Arg Leu Ser Glu Asp
238 130 135 140
240 Arg Val Val Ser Val Ser Phe Arg Val Leu Tyr Pro Ile Val Ile Thr
241 145 150 155 160
243 Ser Leu Gly Val Phe Tyr Asp Ala Ser Asp Val Gly Phe Gln Arg Asn
244 165 170 175
246 Ile Thr Val Lys Leu Tyr Gln Thr Glu Gln Glu Glu Ala Leu Phe Ile
247 180 185 190
249 Ala Arg Phe Ser Pro Pro Ser Cys Gly Val Gln Val Asn Lys Leu Trp
250 195 200 205
252 Tyr Lys Pro Val Glu Gln Phe Ile Leu Pro Glu Ser Phe Glu Gly Thr
253 210 215 220
255 Ile Val Trp Glu Ser Gln Asp Leu His Gly Leu Val Ser Arg Asn Leu
256 225 230 235 240
258 His Arg Val Thr Val Asn Asp Gly Gly Gly Val Leu Arg Val Leu Ala

```

## RAW SEQUENCE LISTING

DATE: 03/16/2006

PATENT APPLICATION: US/10/539,450

TIME: 11:40:04

Input Set : N:\SMITH\PTO.TS.txt

Output Set: N:\CRF4\03162006\J539450.raw

```

259          245          250          255
261 Ala Gly Glu Gly Ala Leu Pro His Glu Phe Met Glu Gly Val Glu Gly
262          260          265          270
264 Val Ala Gly Gly Phe Ile Tyr Thr Val Gln Glu Gly Asp Ala Leu Leu
265          275          280          285
267 Arg Ser Leu Tyr Ser Arg Pro Gln Arg Leu Ala Asp His Ile Gln Asp
268          290          295          300
270 Leu Gln Val Glu Asp Ala Leu Leu Gln Glu Glu Ser Ser Val His Asp
271 305          310          315          320
273 Asp Ile Val Phe Val Asp Val Val Asp Thr Tyr Arg Asn Val Pro Ala
274          325          330          335
276 Lys Leu Leu Asn Phe Tyr Arg Trp Thr Val Glu Ser Thr Ser Phe Asp
277          340          345          350
279 Leu Leu Leu Lys Thr Asp Asp Asp Cys Tyr Ile Asp Leu Glu Ala Val
280          355          360          365
282 Phe Asn Arg Ile Ala Gln Lys Asn Leu Asp Gly Pro Asn Phe Trp Trp
283          370          375          380
285 Gly Asn Phe Arg Leu Asn Trp Ala Val Asp Arg Thr Gly Lys Trp Gln
286 385          390          395          400
288 Glu Leu Glu Tyr Pro Ser Pro Ala Tyr Pro Ala Phe Ala Cys Gly Ser
289          405          410          415
291 Gly Tyr Val Ile Ser Lys Asp Ile Val Asp Trp Leu Ala Gly Asn Ser
292          420          425          430
294 Arg Arg Leu Lys Thr Tyr Gln Gly Glu Asp Val Ser Met Gly Ile Trp
295          435          440          445
297 Met Ala Ala Ile Gly Pro Lys Arg His Gln Asp Ser Leu Trp Leu Cys
298          450          455          460
300 Glu Lys Thr Cys Glu Thr Gly Met Leu Ser Ser Pro Gln Tyr Ser Pro
301 465          470          475          480
303 Glu Glu Leu Ser Lys Leu Trp Glu Leu Lys Glu Leu Cys Gly Asp Pro
304          485          490          495
306 Cys Gln Cys Glu Ala Lys Val Arg
307          500
310 <210> SEQ ID NO: 5
311 <211> LENGTH: 37
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR
318 <400> SEQUENCE: 5
319 cccaagcttg ggctgcaga tcagttggcc ttatttc 37
322 <210> SEQ ID NO: 6
323 <211> LENGTH: 42
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR
330 <400> SEQUENCE: 6
331 aacgcggatc cgcgctgtta tcttgcttga catcgacaag ga 42

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/539,450

DATE: 03/16/2006

TIME: 11:40:05

Input Set : N:\SMITH\PTO.TS.txt

Output Set: N:\CRF4\03162006\J539450.raw

L:18 M:271 C: Current Filing Date differs, Replaced Current Filing Date